Outliers in Cross-Sectional Regressions[‡].

Jesús Mur (*), Jørgen Lauridsen (**),

(*)Department of Economic Analysis.
University of Zaragoza.
Zaragoza. Spain. e-mail: jmur@unizar.es

(**)Corresponding author:
The Econometric Group
Department of Economics.
University of Southern Denmark.
Odense. Denmark. e-mail: jtl@sam.sdu.dk

Abstract

The robustness of the results coming from an econometric application depends to a great extent on the quality of the sampling information. This statement is a general rule that becomes especially relevant in a spatial context where data usually have lots of irregularities.

The purpose of our paper is to examine more closely this question paying attention to one point in particular, namely outliers. The presence of outliers in the sample may be useful, for example in order to break some multicollinearity relations but they may also result in other inconsistencies. The main aspect of our work is that we resolve the discussion in a spatial context, looking closely into the behaviour shown, under several unfavourable conditions, by the most outstanding misspecification tests. For this purpose, we plan and solve a Monte Carlo simulation. The conclusions point to the fact that these statistics react in a different way to the problems posed.

[‡] **Acknowledgements**: This work has been carried out with the financial support of project SEC 2002-02350 of the Ministerio de Ciencia y Tecnología del Reino de España. The authors also wish to thank Ana Angulo for her invaluable and disinterested collaboration.

1. - Introduction

The main purpose of this paper is to examine the relationship between the quality of the sampling information and the trustworthiness of econometric results in a cross-sectional setting. We will focus specifically on outliers.

The existence of outliers is a cause of concern and uncertainty for econometricians. Concern, because it reveals a certain weakness in the data, which always undermines our confidence in the final results. Uncertainty, because any decision, that is taken in these situations, can end up generating undesired effects. Consequently, the attention that the subject has received is not surprising. The works of Hawkins (1980), Chatterjee and Hadi (1988), Belsley (1991) and Barnett and Lewis (1994) are some of the essential references. Furthermore, the literature on the subject has grown in recent years, both in volume and in details, as the use of high frequency series has become more generalised. The works of Chang et al. (1988), Tsay (1988), Perron (1989) or Peña (1990) were pioneer works in this line, which now occupies a preferential position in the modern analysis of time series.

The situation described contrasts notably with that which exists in a context of spatial econometrics. In the first place, because spatial data are prone to the generation of anomalous observations (*heterogeneity*, *instability*, *irregularity*,... very common terms in this area). Nevertheless, the references dedicated explicitly to the subject are scarce (the works of Wartenberg, 1989, and of Haining, 1994 and 1995, are an exception). However, the treatment of outliers is much more mature in the specific area of spatial statistics (Cressie, 1993).

From our point of view, outliers are not necessarily bad, quite the opposite. We essentially agree with Shekhar et al. (2002) when they state that '*Outliers have been informally defined as observations which appear to be inconsistent with the remainder of a set of data, or which deviate so much from other observations so as to arouse suspicions that they were generated by a different mechanism. The identification of outliers can lead to the discovery of unexpected knowledge and has a number of practical applications in (*different*) areas (…)*' (pp. 451-452). These observations become harmful only when they escape from the control of the analyst. Their presence will contaminate the sampling information, distorting the performance of the estimators, etc. If, on the other hand, these observations are conveniently detected and isolated, they may become a source of very valuable information given that they come from bad represented regions of the sample space.

While the framework to be presented naturally serves well as a device to trace the impact of multicollinearity as well as the impact of joint presence of outliers and multicollinearity, the present study concentrates on the outlier problem. An analysis of the effect of multicollinearity in spatial cross-sectional regression appears in Lauridsen and Mur (2004), and an integrative study combining the two anomalities is planned to occur (Lauridsen and Mur, 2005).

The purpose of the present investigation is to address the specific problems caused by outliers when performing misspecification tests in a spatial regression. In section 2 we go deeply into the issues raised by outliers. It is shown that these will have a noticeable impact on most of the statistics and may result in misleading conclusions. A simulation study is carried out in the third section in order to analyse the finite sample impact of anomalous data on tests of misspecification. The paper finishes with a section of conclusions.

2. - Outliers in cross-sectional econometric models

For the moment we will limit ourselves to evaluate the impact that outliers have on the misspecification statistics habitually used in cross-sectional econometric models**,** that is, on Moran's I, LM-ERR, LM-EL and KR, which address the problem of spatial dependence in the error term, together with the LM-LAG and the LM-LE whose objective is to analyse the dynamic structure of the equation. To these we add the SARMA test, whose null hypothesis is composite (static structure in the equation and a white noise error term). In Appendix 1 there is a brief presentation. With respect to our work, it is important to point out that the seven tests are constructed from the residuals of the LS estimation. Given that these residuals react in a different way to the presence of anomalies in the sample, this sensitivity should appear, at least in part, also in the tests. Next we discuss both topics one by one.

The impact of these points depends, in the first place, on the dimensions of the anomaly itself, to which must be added, in this case, its geographical location and the structure of cross-sectional dependencies that exist in the data. To appreciate this more clearly, we can proceed with the following model which incorporates an outlier:

$$\left.\begin{array}{l} y = X\beta + \pi d + u \\ d_s = \begin{cases} 0 & s \neq r \\ 1 & s = r \end{cases}; \quad u \sim N(0;\sigma^2) \end{array}\right\} \qquad (1)$$

The anomaly consists in a shift in the mean of observation r and its value is π. If we omit this fact to propose a linear equation: $y = X\beta + v$, we will have committed a specification error of not including a relevant variable in the systematic part. The consequences derived from this error are well known: the LS estimators are biased and the distribution of the residuals is no longer centred on zero:

$$\left.\begin{array}{l} y = X\beta + \pi d + u = X\beta + v \\ v = \pi d + u \to \begin{cases} E[v] = \pi d \\ V[v] = \sigma^2 I \end{cases} \end{array}\right\} \Rightarrow \begin{cases} \hat{\beta} = [X'X]^{-1}X'y \to \begin{cases} E[\hat{\beta}] = \beta + \pi[X'X]^{-1}X'd \\ V[\hat{\beta}] = \sigma^2[X'X]^{-1} \end{cases} \\ \hat{v} = y - X\hat{\beta} = Mv \to \begin{cases} E[\hat{v}] = \pi Md = m_r \\ V[\hat{v}] = \sigma^2 M \end{cases} \end{cases} \quad (2)$$

where M is the matrix $\left[I - X(X'X)^{-1}X'\right]$; vector $Md=m_r$ corresponds to the r-th column of matrix M. The special feature that we wish to introduce is either an error term with spatial dependence (SAR or SMA), or the presence of dynamic elements in the main equation of (1).

In the case of residual dependence with an SAR structure, the equations are:

$$\left.\begin{array}{l} y = X\beta + v \\ v = \rho Wv + u \\ u = \varepsilon + \pi d; \varepsilon \sim iidN(0,\sigma^2) \end{array}\right\} \Rightarrow \begin{cases} y = X\beta + v \\ v = [I-\rho W]^{-1}u = v^* + \pi d^*_{SAR} \\ v^* = [I-\rho W]^{-1}\varepsilon; \;\; d^*_{SAR} = [I-\rho W]^{-1}d \end{cases} \quad (3)$$

Unlike what occurs in case (2), the shift in the mean will affect all the error terms to a different extent (πd^*_{SAR}; that is the r-th column of the matrix $[I-\rho W]^{-1}$ multiplied by π). The error of omitting the outlier in observation r will give biased LS estimators:

$$\hat{\beta} = [X'X]^{-1}X'y \Rightarrow E[\hat{\beta}] = \beta + \pi[X'X]^{-1}X'd^*_{SAR} \quad (4)$$

The LS residuals will not be centred on zero:

$$\hat{v} = y - X\hat{\beta} = Mv \to \begin{cases} E[\hat{v}] = \pi Md^*_{SAR} \\ V[\hat{v}] = \sigma^2 M[I-\rho W]^{-2}M \end{cases} \quad (5)$$

If a moving average structure dominates the error term, the results are similar. The LS estimators are biased and the vector of residuals also undergoes a shift:

$$\hat{\beta}=[X'X]^{-1}X'y \Rightarrow E\left[\hat{\beta}\right]=\beta+\pi[X'X]^{-1}X'd^{*}_{SMA} \qquad (6)$$

$$\hat{v}=y-X\hat{\beta}=Mv \rightarrow \begin{cases} E[\hat{v}]=\pi M d^{*}_{SMA} \\ V[\hat{v}]=\sigma^2 M[I-\theta W]^2 M \end{cases} \qquad (7)$$

where ρ is the parameter of the process and $d^{*}_{SMA}=[I-\rho W]d$. The impact of the outlier is spatially concentrated in this case, because the vector d^{*}_{SMA} corresponds to the r-th column of the SMA matrix $[I-\rho W]$. Lastly, in the case of substantive spatial dependence, the impact of the outlier reaches to all the observations:

$$\left.\begin{array}{l} y=\rho Wy+X\beta+v \\ v=\varepsilon+\pi d;\varepsilon \sim iidN(0,\sigma^2) \end{array}\right\} \Rightarrow \begin{cases} y=[I-\rho W]^{-1}(X\beta+\pi d)+[I-\rho W]^{-1}\varepsilon= \\ \quad =[I-\rho W]^{-1}X\beta+\pi d^{*}_{SAR}+[I-\rho W]^{-1}\varepsilon \end{cases} \qquad (8)$$

The bias of the LS estimators is strong, as is the shift in the first order moment of the residuals:

$$E\left[\hat{\beta}\right]=[X'X]^{-1}X'(I-\rho W)^{-1}X\beta+\pi d^{*}_{SAR} \qquad (9)$$

$$\hat{v}=y-X\hat{\beta}=My \rightarrow \begin{cases} E[\hat{v}]=M(I-\rho W)^{-1}X\beta+\pi M d^{*}_{SAR} \\ V[\hat{v}]=\sigma^2 M[I-\rho W]^{-2} M \end{cases} \qquad (10)$$

In spite of the above results, it is not easy to define the impact of these errors on the misspecification tests. Moran's I suffers a shift, of an indeterminate sign and quantity, that affects the first order moment[1]. Now, the new expected value, under the null hypothesis of no correlation in the error term and assuming an anomaly similar to that in (1), becomes:

$$E[I]_{H_0}=\left(\frac{R}{S_0}\right)\left(\frac{\sigma^2 trMW+\pi^2 m_r' W m_r}{\sigma^2(R-k)+\pi^2(1-p_{rr})}\right) \qquad (11)$$

where p_{rr} is the r-th element of the main diagonal of the matrix $P=X(X'X)^{-1}X'$. This element measures the leverage of the associated observation and is bounded as: $0 \le p_{rr} \le 1$, so that it will

---

[1] $E[I]_{H_0}=\left(\frac{R}{S_0}\right)\left(\frac{trMW}{R-k}\right)$ in normal circumstances.

always be true that $\pi^2(1-p_{rr}) \geq 0$. However, the impact on the numerator is uncertain because the quadratic form $m_r^{'}Wm_r$ is undefined. The consequences on the second order moments of the probability distribution of Moran's I are even vaguer.

The Lagrange Multipliers also undergo several adjustments. For example, the ML estimation of the model of (3), introducing an explicit error of not detecting an outlier in the r-th observation, leads to the score vector:

$$g(\gamma) = \frac{\partial l}{\partial \gamma} = \begin{bmatrix} \frac{\partial l}{\partial \beta} \\ \frac{\partial l}{\partial \rho} \\ \frac{\partial l}{\partial \sigma^2} \end{bmatrix} = \frac{1}{\sigma^2}\begin{bmatrix} X'B'(\pi d + \varepsilon) \\ (\pi d + \varepsilon)'B^{-1}W(\pi d + \varepsilon) - trWB^{-1} \\ -\frac{R}{2} + \frac{(\pi d + \varepsilon)'(\pi d + \varepsilon)}{2\sigma^2} \end{bmatrix} \tag{12}$$

where *l* is the log-likelihood, γ the vector [β,ρ,σ^2]’ and B the matrix [I-ρW]. Its expected value under the null hypothesis of no correlation (H_0: ρ=0) will not be zero:

$$E\left[g(\gamma)\right]_{|H_0:\rho=0} = \frac{\pi}{\sigma^2}\begin{bmatrix} X'd \\ 0 \\ \frac{\pi}{2\sigma^2} \end{bmatrix} = \frac{\pi}{\sigma^2}\begin{bmatrix} x_r^{'} \\ 0 \\ \frac{\pi}{2\sigma^2} \end{bmatrix} \tag{13}$$

where x_r is the (1xk) vector corresponding to the exogenous variables in the contaminated point. The outlier, as can be seen, does not have any influence on the expected value of the second element of the score, corresponding to the ML estimation of ρ, that is still zero. Furthermore, the Hessian matrix, again under the null, is no longer block diagonal:

$$H(\gamma) = -E\left[\frac{\partial^2 l}{\partial\gamma\partial\gamma'}\right]_{|H_0:\rho=0} = \frac{1}{\sigma^2}\begin{bmatrix} (X'X) & 2\pi X'Wd & \frac{\pi}{\sigma^2}x_r^{'} \\ 2\pi d'W'X & 2\sigma^2 S_0 + \pi^2 s_r & 0 \\ \frac{\pi}{\sigma^2}x_r & 0 & \frac{R}{2\sigma^2}\left(1+\frac{2}{R}\frac{\pi^2}{\sigma^2}\right) \end{bmatrix} \tag{14}$$

where s_r is the r-th element of the main diagonal of W’W (if this matrix is of a binary type, s_r corresponds to the number of contacts of the contaminated region).

These results will be unknown to the analyst who limits himself to working as usual. Particularly, to test the null hypothesis that the error term is a white noise (H_0: $\rho = 0$) against the alternative of spatial dependence (H_A:$\rho \neq 0$), the LM-ERR specified will be:

$$\mathrm{LM{-}ERR}^{A} = [g(\hat{\gamma})]'\left[H^{A}(\hat{\gamma})\right]^{-1}[g(\hat{\gamma})] \tag{15}$$

The superscript A indicates that we are dealing with an approximation to the relevant element. Really, the structure of the Lagrange Multiplier that should have been used is the following:

$$\mathrm{LM-ERR} = [g(\hat{\gamma}) - Eg(\hat{\gamma})]'[H(\hat{\gamma})]^{-1}[g(\hat{\gamma}) - Eg(\hat{\gamma})] \underset{as}{\sim} \chi^2(1) \tag{16}$$

To relate both statistics (LM-ERR and LM-ERR^A) we may decompose the Hessian matrix of (14) into:

$$H(\gamma) = H^{A}(\gamma) + H^{B}(\gamma) = \frac{1}{\sigma^2}\begin{bmatrix} (X'X) & 0 & 0 \\ 0 & 2\sigma^2 S_0 & 0 \\ 0 & 0 & \dfrac{R}{2\sigma^2} \end{bmatrix} + \frac{1}{\sigma^2}\begin{bmatrix} 0 & 2\pi X'Wd & \dfrac{\pi}{\sigma^2}x_r' \\ 2\pi d'W'X & \pi^2\sigma^2 s_r & 0 \\ \dfrac{\pi}{\sigma^2}x_r & 0 & \dfrac{\pi^2}{\sigma^4} \end{bmatrix} \tag{17}$$

Introducing the last expression into (16) we get:

$$\begin{aligned} \mathrm{LM-ERR} &= [g(\hat{\gamma}) - Eg(\hat{\gamma})]'[H(\hat{\gamma})]^{-1}[g(\hat{\gamma}) - Eg(\hat{\gamma})] = \\ &= [g(\hat{\gamma}) - Eg(\hat{\gamma})]'\left[\left(H^{A}(\hat{\gamma})\right)^{-1} - H^{*}(\hat{\gamma})\right][g(\hat{\gamma}) - Eg(\hat{\gamma})] = \\ &= g(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1}g(\hat{\gamma}) - Eg(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1}Eg(\hat{\gamma}) - [g(\hat{\gamma}) - Eg(\hat{\gamma})]'H^{*}(\hat{\gamma})[g(\hat{\gamma}) - Eg(\hat{\gamma})] = \\ &= g(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1}g(\hat{\gamma}) - \mathit{lmerr}_1 - \mathit{lmerr}_2 \end{aligned} \tag{18}$$

where $H^{*}(\hat{\gamma}) = \left(H^{A}(\hat{\gamma})\right)^{-1}\left[\left(H^{A}(\hat{\gamma})\right)^{-1} + \left(H^{B}(\hat{\gamma})\right)^{-1}\right]^{-1}\left(H^{A}(\hat{\gamma})\right)^{-1}$. In short, the relationship we are looking for is:

$$\mathrm{LM{-}ERR}^{A} = g(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1}g(\hat{\gamma}) = \mathrm{LM-ERR} + \mathit{lmerr}_1 + \mathit{lmerr}_2 \tag{19}$$

The term *lmerr*$_1$ tends, with R, to a positive constant:

$$\underset{R\to\infty}{p\lim}\left\{Eg(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1}Eg(\hat{\gamma})\right\} = \underset{R\to\infty}{p\lim}\left\{\pi^2\left(p_{rr}+\frac{\pi^2}{2R\sigma^2}\right)\right\} = p_{rr}\pi^2 \geq 0 \qquad (20)$$

On the other hand, the term *lmerr₂* is a quadratic form of a random vector, which obeys a Central Limit Theorem, $\sqrt{R}\left[g(\gamma)-Eg(\gamma)\right]\xrightarrow{D}N\left[0;H(\gamma)^{-1}\right]$, in a (now) known matrix, $H^{*}(\hat{\gamma})$. The probability distribution of this quadratic form is not standard, and corresponds to that of a sum of mutually independent chi-squared variables, with weights λ_j:

$$lmerr_2 = [g(\hat{\gamma})-Eg(\hat{\gamma})]'H^{*}(\hat{\gamma})[g(\hat{\gamma})-Eg(\hat{\gamma})] = \sum_{j=1}^{R}\lambda_j z_j^2 \qquad (21)$$

The random variables z_j are independent unit normal N(0,1) variables and the elements $\{\lambda_j$, j=1, 2, ...R$\}$ are the eigenvalues of the matrix $L(\gamma)'H^{*}(\gamma)L(\gamma)$, where $L(\gamma)$ is the matrix that factorizes the information matrix of (14): $H(\gamma)=L(\gamma)L(\gamma)'$ (Kendall and Stuart, 1977).

The consequences on LM-LAG test can be followed using the model of expression (8). The score vector corresponding to this specification will be:

$$g(\gamma)=\frac{\partial l}{\partial\gamma}=\begin{bmatrix}\frac{\partial l}{\partial\beta}\\ \frac{\partial l}{\partial\rho}\\ \frac{\partial l}{\partial\sigma^2}\end{bmatrix}=\frac{1}{\sigma^2}\begin{bmatrix}X'[\pi d+\varepsilon]\\ y'W'(\pi d+\varepsilon)-trWB^{-1}\\ -\frac{R}{2}+\frac{(\pi d+\varepsilon)'(\pi d+\varepsilon)}{2\sigma^2}\end{bmatrix} \qquad (22)$$

It should be remembered that, now, ρ is the parameter that accompanies the lag of the endogenous variable in the main equation of the model; B continues to be the matrix [I-ρW]. The expected value of the score of (22), under the hypothesis of static structure (H_0: ρ=0), is again different from zero:

$$E\left[g(\gamma)\right]_{|H_0:\rho=0}=\frac{\pi}{\sigma^2}\begin{bmatrix}x_r^{'}\\ \beta'X'Wd\\ \frac{\pi}{2\sigma^2}\end{bmatrix} \qquad (23)$$

The information matrix, under the null, is of the general type:

$$H(\gamma) = -E\left[\frac{\partial^2 l}{\partial\gamma\partial\gamma'}\right]_{|H_0:\rho=0} =$$

$$= \frac{1}{\sigma^2}\begin{bmatrix} (X'X) & X'W(X\beta+\pi d) & \frac{\pi}{\sigma^2}x_r' \\ (X\beta+\pi d)'W'X & \begin{Bmatrix} \beta'X'W^2(X\beta+2\pi d)+ \\ +\pi^2 s_r + 2\sigma^2 S_0 \end{Bmatrix} & \frac{\pi}{\sigma^2}\beta'X'Wd \\ \frac{\pi}{\sigma^2}x_r & \frac{\pi}{\sigma^2}d'W'X\beta & \frac{R}{2\sigma^2}\left(1+\frac{2}{R}\frac{\pi^2}{\sigma^2}\right) \end{bmatrix} \qquad (24)$$

with W^2=W'W. This matrix admits a decomposition similar to that used before, so that:

$$H(\gamma) = H^A(\gamma) + H^B(\gamma) =$$

$$\frac{1}{\sigma^2}\left\{\begin{bmatrix} (X'X) & X'WX\beta & 0 \\ \beta'X'W'X & \begin{Bmatrix} \beta'X'W^2X\beta+ \\ 2\sigma^2 S_0 \end{Bmatrix} & 0 \\ 0 & 0 & \frac{R}{2\sigma^2} \end{bmatrix} + \begin{bmatrix} 0 & \pi X'Wd & \frac{\pi}{\sigma^2}x_r' \\ \pi d'W'X & \begin{Bmatrix} 2\pi\beta'X'W^2 d+ \\ \pi^2 s_r \end{Bmatrix} & \frac{\pi}{\sigma^2}\beta'X'Wd \\ \frac{\pi}{\sigma^2}x_r & \frac{\pi}{\sigma^2}d'W'X\beta & \frac{\pi^2}{\sigma^4} \end{bmatrix}\right\}$$

(25)

The last result allows us to develop the LM-LAG test as:

$$LM-LAG = [g(\hat{\gamma}) - Eg(\hat{\gamma})]'[H(\hat{\gamma})]^{-1}[g(\hat{\gamma}) - Eg(\hat{\gamma})] =$$

$$= g(\hat{\gamma})'\left(H^A(\hat{\gamma})\right)^{-1}g(\hat{\gamma}) - Eg(\hat{\gamma})'\left(H^A(\hat{\gamma})\right)^{-1}Eg(\hat{\gamma}) - [g(\hat{\gamma}) - Eg(\hat{\gamma})]'H^*(\hat{\gamma})[g(\hat{\gamma}) - Eg(\hat{\gamma})] =$$

$$= g(\hat{\gamma})'\left(H^A(\hat{\gamma})\right)^{-1}g(\hat{\gamma}) - lmlag_1 - lmlag_2$$

(26)

The matrix H*(γ) maintains, formally, the same structure as that already used, it being sufficient to update the contents of the constituent matrices $\left(H^A(\hat{\gamma})\right)$ and $\left(H^B(\hat{\gamma})\right)$. We can write:

$$LM-LAG^A = g(\hat{\gamma})'\left(H^A(\hat{\gamma})\right)^{-1}g(\hat{\gamma}) = LM-LAG + lmlag_1 + lmlag_2 \qquad (27)$$

The term $lmlag_1$ tends, with R, to a positive constant:

$$\underset{R\to\infty}{p\lim}\left\{Eg(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1}Eg(\hat{\gamma})\right\} =$$

$$\underset{R\to\infty}{p\lim}\left\{\frac{\pi^2}{\sigma^4}\left(\sigma^2 p_{rr} + V[\rho]\left[(d-p_r)'WX\beta\right]^2 + \frac{\pi^2}{2R}\right)\right\} = \frac{\pi^2}{\sigma^4}\left(\sigma^2 p_{rr} + V[\rho]\left[(d-p_r)'WX\beta\right]^2\right) \tag{28}$$

where V[ρ] is the probability limit of the variance of the ML estimator of parameter ρ: $V[\rho] = \left[\frac{\beta'X'W'MWX\beta}{\sigma^2}\right]^{-1}$ and p_r the r-th column of matrix P. The behaviour of the term $lmlag_2$ is more imprecise, although the result of (21) can be adapted here:

$$lmlag_2 = [g(\hat{\gamma}) - Eg(\hat{\gamma})]'H^*(\hat{\gamma})[g(\hat{\gamma}) - Eg(\hat{\gamma})] = \sum_{j=1}^{R}\lambda_j z_j^2 \tag{29}$$

The variables z_j are still independent unit normal N(0,1) variables and the weights {λ_j, j=1, 2, ...R} are the eigenvalues of the matrix $L(\gamma)'H^*(\gamma)L(\gamma)$, where $L(\gamma)$ is the matrix that factorizes the information matrix of expression (24).

To conclude this discussion it remains to consider the case of the SARMA test. Its null hypothesis is composite, so that the model must be general:

$$\left.\begin{array}{l} y = \rho Wy + X\beta + v \\ v = \theta Wv + u \\ u = \varepsilon + \pi d; \varepsilon \sim iidN(0,\sigma^2) \end{array}\right\} \tag{30}$$

There are no surprises in the score:

$$g(\gamma)=\frac{\partial l}{\partial \gamma}=\begin{bmatrix}\frac{\partial l}{\partial \beta}\\ \frac{\partial l}{\partial \rho}\\ \frac{\partial l}{\partial \theta}\\ \frac{\partial l}{\partial \sigma^2}\end{bmatrix}=\frac{1}{\sigma^2}\begin{bmatrix}X'D'[\pi d+\varepsilon]\\ y'WD'(\pi d+\varepsilon)-trWB^{-1}\\ (\pi d+\varepsilon)'D^{-1}W(\pi d+\varepsilon)-trWD^{-1}\\ -\frac{R}{2}+\frac{(\pi d+\varepsilon)'(\pi d+\varepsilon)}{2\sigma^2}\end{bmatrix}\Rightarrow \tag{31}$$

$$\Rightarrow E\left[g(\gamma)\right]_{|H_0:\rho=\theta=0}=\frac{\pi}{\sigma^2}\begin{bmatrix}x_r'\\ \beta'X'Wd\\ 0\\ \frac{\pi}{2\sigma^2}\end{bmatrix}$$

where B=[I-ρW] and D=[I-θW]. Again the presence of the outlier does not affect the term of the score associated with the ML estimator of θ, whose expected value is zero. The structure of the Hessian matrix associated with this case is complex even under the null of no spatial interaction:

$$H(\gamma)=-E\left[\frac{\partial^2 l}{\partial\gamma\partial\gamma'}\right]_{|H_0:\rho=\theta=0}=$$

$$=\frac{1}{\sigma^2}\begin{bmatrix}(X'X) & X'W(X\beta+\pi d) & 2\pi X'Wd & \frac{\pi}{\sigma^2}x_r' \\ (X\beta+\pi d)'W'X & \begin{Bmatrix}\beta'X'W^2(X\beta+2\pi d)+\\ +\pi^2 s_r+2\sigma^2 S_0\end{Bmatrix} & \begin{Bmatrix}2(\pi d'W^2X\beta-\\ \pi^2 s_r+\sigma^2 S_0)\end{Bmatrix} & \frac{1}{\sigma^2}2\pi d'W'X\beta \\ 2\pi d'W'X & \begin{Bmatrix}2(\pi d'W^2X\beta-\\ \pi^2 s_r+\sigma^2 S_0)\end{Bmatrix} & 2\sigma^2 S_0+\pi^2 s_r & 0 \\ \frac{\pi}{\sigma^2}x_r & \frac{1}{\sigma^2}2\pi\beta'X'Wd & 0 & \frac{R}{2\sigma^2}\left(1+\frac{2}{R}\frac{\pi^2}{\sigma^2}\right)\end{bmatrix} \tag{32}$$

In any case, that matrix can be decomposed into the sum of two matrices, as has already been done on other occasions:

$$H(\gamma) = H^{A}(\gamma) + H^{B}(\gamma) =$$

$$\frac{1}{\sigma^2}\left\{\begin{bmatrix} X'X & X'WX\beta & 0 & 0 \\ \beta'X'W'X & \left\{\begin{matrix}\beta'X'W^2X\beta + \\ 2\sigma^2 S_0\end{matrix}\right\} & 2S_0 & 0 \\ 0 & 2S_0 & 2S_0 & 0 \\ 0 & 0 & 0 & \dfrac{R}{2\sigma^2} \end{bmatrix} + \\ + \begin{bmatrix} 0 & \pi X'Wd & 2\pi X'Wd & \dfrac{\pi}{\sigma^2} x_r' \\ \pi d'W'X & \left\{\begin{matrix}2\pi\beta'X'W^2 d + \\ \pi^2 s_r\end{matrix}\right\} & \left\{\begin{matrix}2\pi\beta'X'W^2 d - \\ -2\pi^2 s_r\end{matrix}\right\} & \dfrac{2\pi}{\sigma^2}\beta'X'Wd \\ 2\pi d'W'X & \left\{\begin{matrix}2\pi\beta'X'W^2 d - \\ -2\pi^2 s_r\end{matrix}\right\} & \pi^2 s_r & 0 \\ \dfrac{\pi}{\sigma^2} x_r & \dfrac{2\pi}{\sigma^2} d'W'X\beta & 0 & \dfrac{\pi^2}{\sigma^4} \end{bmatrix}\right\} \qquad (33)$$

The SARMA statistic can finally be expressed as:

$$\mathrm{SARMA} = [g(\hat{\gamma}) - Eg(\hat{\gamma})]'[H(\hat{\gamma})]^{-1}[g(\hat{\gamma}) - Eg(\hat{\gamma})] =$$
$$= g(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1} g(\hat{\gamma}) - Eg(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1} Eg(\hat{\gamma}) - [g(\hat{\gamma}) - Eg(\hat{\gamma})]' H^{*}(\hat{\gamma})[g(\hat{\gamma}) - Eg(\hat{\gamma})] =$$
$$= g(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1} g(\hat{\gamma}) - \mathit{sarma}_1 - \mathit{sarma}_2$$

(34)

The structure of the matrix H*(γ), and that of the matrices $H^{A}(\hat{\gamma})$ and $H^{B}(\hat{\gamma})$ that form it, remain the same. The SARMA statistic, wrongly specified because it does not reflect the anomaly of the r-th observation, corresponds to:

$$\mathrm{SARMA}^{A} = g(\hat{\gamma})'\left(H^{A}(\hat{\gamma})\right)^{-1} g(\hat{\gamma}) = \mathrm{SARMA} + \mathit{sarma}_1 + \mathit{sarma}_2 \qquad (35)$$

The term *sarma*$_1$ has the same probability limit as that corresponding to *lmlag*$_1$:

$$\underset{R\to\infty}{\mathrm{p}\lim}\, sarma_1 = \underset{R\to\infty}{\mathrm{p}\lim}\, lmlag_1 = \underset{R\to\infty}{\mathrm{p}\lim}\left\{Eg(\hat{\gamma})'\left(H^A(\hat{\gamma})\right)^{-1}Eg(\hat{\gamma})\right\} =$$

$$\underset{R\to\infty}{\mathrm{p}\lim}\left\{\frac{\pi^2}{\sigma^4}\left(\sigma^2 p_{rr} + V[\rho]\left[\left(d-p_r\right)'WX\beta\right]^2 + \frac{\pi^2}{2R}\right)\right\} = \frac{\pi^2}{\sigma^4}\left(\sigma^2 p_{rr} + V[\rho]\left[\left(d-p_r\right)'WX\beta\right]^2\right) \tag{36}$$

The term $sarma_2$ is again a quadratic form of a vector of standardised normal variables on a diagonal matrix with elements $\{\lambda_j$, j=1, 2, ...R$\}$, so that:

$$sarma_2 = [g(\hat{\gamma}) - Eg(\hat{\gamma})]'H^*(\hat{\gamma})[g(\hat{\gamma}) - Eg(\hat{\gamma})] = \sum_{j=1}^{R}\lambda_j z_j^2 \tag{37}$$

These weights $\{\lambda_j$, j=1, 2, ...R$\}$ coincide with the eigenvalues of the matrix $L(\gamma)'H^*(\gamma)L(\gamma)$, where $L(\gamma)$ is the matrix that factorizes the information matrix of (32).

3. - Monte Carlo results

In the previous section, the impact of having low quality sampling information on the misspecification statistics has been discussed. That results allow us to affirm that the presence of outliers hinder the performance of these statistics. Particularly, it has been shown that the size and, surely, also the power of the tests are affected by a downwards bias.

Two models have been simulated: a static one with a structure of residual dependence and the other, dynamic with a white noise error term. The model simulated in the static case has a simple design:

$$y_r = \beta_0 + \beta_1 x_{1r} + v_r; \quad r = 1,2,....,R \tag{38}$$

The error term v_t responds to the different cases that have been planned, SAR or SMA, including outliers:

SAR STRUCTURE **SMA STRUCTURE**

$$\begin{cases} y = X\beta + v \\ v = \rho W v + u \\ u = \varepsilon + \pi d; \varepsilon \sim iidN(0,\sigma_\varepsilon^2) \end{cases} \qquad \begin{cases} y = X\beta + v \\ v = u - \rho W u \\ u = \varepsilon + \pi d; \varepsilon \sim iidN(0,\sigma_\varepsilon^2) \end{cases} \tag{39}$$

where σ_ε^2 has been set equal to 1. The *d* term stands for the dummy variable used in (1) and π it is a parameter that defines the outlier. The parameters β_j (j=0,1,2) where set to 10. Furthermore, three systems of regions, always of a regular shape, and of sizes 25, 100 and 225, have been introduced as represented in Appendix 2. The matrix W has been specified accordingly, row-standardised. Only positive values of parameter ρ, ranging between 0 and 0.99 have been simulated and the number of simulation is 1000 for each case.

Different combinations of location and value of the outlier have been simulated. For the moment, we have only developed two possible localizations for the anomaly, peripheral or central. The former means intervening in the regions 25, 100 or 225, depending on the system simulated, while in the latter the observations 13, 45 or 112 have been altered. Lastly, the value of the outlier depends on the dispersion of the vector ε obtained previously. Concretely, if we denote the standard deviation of this vector by $\hat{\sigma}_\varepsilon$, parameter π has been set equal to 0 (there is no anomaly), $2.5\hat{\sigma}_\varepsilon$ (the anomaly is small), $5\hat{\sigma}_\varepsilon$ (the anomaly is relevant) or $7.5\hat{\sigma}_\varepsilon$ (the anomaly is very high). For ease of exposition, we only present the results corresponding to the case of a big outlier (π equal to $7.5\hat{\sigma}_\varepsilon$) and a central position.

Finally, the dynamic model simulated can be expressed as:

$$\begin{cases} y = \rho Wy + X\beta + v \\ u = \varepsilon + \pi d; \varepsilon \sim iidN(0, \sigma_\varepsilon^2) \end{cases} \tag{40}$$

The specifications about the different elements that intervene have been maintained in this case.

The principal results obtained are gathered in Table 1, dedicated to the empirical size, and in Figures 1 to 7 which represent the estimated power functions of the tests.

TABLE 1: Empirical size of the tests for a theoretical significance level of 0.05.(*)

	R=25		**R=100**		**R=225**	
	No Outliers	1 Outlier	No Outliers	1 Outlier	No Outliers	1 Outlier
Moran's I	0.034	0.009	0.043	0.029	0.050	0.047
LM-ERR	0.044	0.014	0.049	0.030	0.048	0.049
LM-LAG	0.068	0.033	0.047	0.025	0.061	0.061
LM-EL	0.044	0.027	0.054	0.024	0.052	0.050
LM-LE	0.070	0.017	0.054	0.024	0.063	0.065
SARMA	0.053	0.007	0.048	0.021	0.052	0.057
KR	0.061	0.033	0.050	0.035	0.050	0.051

(*) 95% Confidence interval for *p* (probability of rejecting the null hypothesis), around the theoretical value for the case of 1000 replications, is $0.036 < p < 0.064$.

The results reflected in Table 1 do not contain surprises. As we said, the presence of an outlier in the sample has harmful effects on the size of the statistics, especially when the sample is small and the outlier is big. Its influence is evident in samples of 25 observations where the empirical size frequently falls below 1.0% due to the presence of the outlier. The KR statistic seems to be the most robust to their presence, while the different Lagrange multipliers show a greater sensitivity, sharpened in the Moran's I. The situation improves slightly using 100 observations, the empirical size is reduced to roughly a half from the theoretical one of 5.0%, and the impact of the outlier goes unnoticed in a sample made of 225 data.

The results corresponding to the estimated power functions are presented in graphic form in Figures 1 to 7 (the details can be obtained from the authors upon request).

Some aspects of these graphs were already well known beforehand. The weakness of the spatial dependency tests in a context of small samples is one of them. Another is the worsening of the performance of the tests when a moving average is used in the alternative hypothesis. Neither is the greater reliability of Moran's I for detecting dependency processes in the error term nor the weaknesses of the KR test, especially in SMA structures (Florax and de Graaff, 2004), anything new. The figures also illustrate the low level of specificity of the traditional statistics which react to all types of spatial dependence. This is the *raison d'être* of the robust Lagrange multipliers (LM-EL and LM-LE).

Our results ratify the non-neutrality of outliers. The impact appears to be, in all the cases, a loss of power. The most significant falls occur when the sample is small and when the process simulated in the alternative hypothesis is an SMA. In such circumstances, the KR and LM-EL tests do not reach the minimum percentage of rejections of 50% even using autocorrelation coefficients higher than 0.90 (see Figures 3 and 4). Furthermore, the KR test appears to be the most robust to the presence of outliers when the dependency structure is an SAR (no matter residual or substantive).

The distortions created by outliers are less important when there is substantive dependence in the equation. In this case, as can be seen in Figures 5, 6 and 7 the power cuts suffered by the tests are reduced. The LM-LAG test is the one that performs best in this situation, although its superiority weakens as the sample size increases. With a sample of 225 observations, the differences between this test and the LM-LE (robust to misspecifications of the alternative hypothesis) are inappreciable.

Lastly, another aspect, that we wish to underline, is that, although it seems evident that the incidence of the outliers is reduced as the sample size increases, certain dysfunctions in the performing of these tests are still observable when using a sample of 225 observations. It should be borne in mind that only one observation has been contaminated in spite of increasing the sample size, which does not seem very reasonable in real circumstances: as the number of observations increases, the probability of there being a larger number of outliers also grows. This possibility goes beyond the objectives of the present paper, although it forms part of the agenda for future investigation.

4. - Conclusions and final remarks

The objective of this paper was to examine the influence of atypical observations on the misspecification tests more often used in a context of spatial econometric modelling.

The analytical approximation resolved in the second section allows us to affirm that the probability distributions corresponding to the statistics used are influenced by the impact of outliers. These distributions tend, in the first case, to be displaced to the right of the probability space. The worst situation combines a large outlier and a high leverage in the exogenous variables. This outline corresponds to an observation that is anomalous both in the space spanned by X and in vector y. A high leverage implies that the observation tends to take control of the regression forcing it to pass through its neighbourhood. As a consequence, the associated residual will be negligible (first error: the residual should be important enough to warn of the anomaly) at the expense of distorting the informational content supplied by the other observations (second error: the net of spatial dependencies is distorted).

The simulation carried out in the third section has served to corroborate some hypothesis. It was foreseeable that the impact of the outliers would maintain an inverse relationship with the sample size. Neither is it surprising that the same occurs with respect to the parameter of spatial dependence. However, unexpected results have also been observed. For example, the fact that the anomalies in the data have less incidence, when they arise from a process of substantive spatial dependence. The sensibilities shown by the different statistics is another aspect to highlight, including the apparently robust behaviour of the KR statistic.

Finally, we wish to insist that this paper is nothing more than a first approximation to the problem of outliers in cross-sectional econometric models. We have analysed a limited number of combinations with which we have been reaching a few conclusions. Nevertheless, the cases that remain to be studied (increasing number of outliers, different geographical locations, characteristics of the variables, joint impact of multicollinearity and outliers, etc) provide even more interesting topics which merit further investigation.

APPENDIX 1: Misspecification tests used.

The tests used always refer to the model of the null hypothesis; that is, of the static type such as: y = Xβ + u. This model has been estimated by LS, where $\hat{\sigma}^2$ and $\hat{\beta}$ are the corresponding LS estimations and $\hat{u}$ the residual series. The tests are the following (see Anselin and Florax, 1995, or Florax and de Graaff, 2004, for details):

Moran's I Test: $$I = \frac{R}{S_0}\frac{\hat{u}'W\hat{u}}{\hat{u}'\hat{u}}; \qquad S_0 = \sum_{r=1}^{R}\sum_{s=1}^{R} w_{rs} \quad \text{(A.1)}$$

LM-ERR Test: $$LM - ERR = \left(\frac{\hat{u}'W\hat{u}}{\hat{\sigma}^2}\right)^2 \frac{1}{T_1}; \; T_1 = tr\left[W'W + WW\right] \quad \text{(A.2)}$$

LM-EL Test: $$LM - EL = \frac{\left(\frac{\hat{u}'W\hat{u}}{\hat{\sigma}^2} - \frac{T_1}{R\hat{J}_{\rho\text{-}\beta}}\frac{\hat{u}'Wy}{\hat{\sigma}^2}\right)^2}{T_1 - \frac{T_1^2}{R\hat{J}_{\rho\text{-}\beta}}} \quad \text{(A.3)}$$

KR Test: $$KR = h_R \frac{\hat{\gamma}'Z'Z\hat{\gamma}}{\hat{e}'\hat{e}} \quad \text{(A.4)}$$

LM-LAG Test: $$LM - LAG = \frac{1}{R\hat{J}_{\rho-\beta}}\left(\frac{\hat{u}'Wy}{\hat{\sigma}^2}\right)^2 \quad \text{(A.5)}$$

LM-LE Test: $$LM - LE = \frac{\left(\frac{\hat{u}'Wy}{\hat{\sigma}^2} - \frac{\hat{u}'W\hat{u}}{\hat{\sigma}^2}\right)^2}{R\hat{J}_{\rho\text{-}\beta} - T_1} \quad \text{(A.6)}$$

SARMA Test: $$SARMA = \frac{\left(\frac{\hat{u}'Wy}{\hat{\sigma}^2} - \frac{\hat{u}'W\hat{u}}{\hat{\sigma}^2}\right)^2}{R\hat{J}_{\rho\text{-}\beta} - T_1} + \frac{1}{T_1}\left(\frac{\hat{u}'W\hat{u}}{\hat{\sigma}^2}\right)^2 \quad \text{(A.7)}$$

Moreover, $R\hat{J}_{\rho-\beta} = T_1 + (\hat{\beta}'X'WMWX\hat{\beta})/\hat{\sigma}^2$ and $M=[I-X(X'X)^{-1}X']$. Furthermore, $\hat{e}$ is the vector of residuals from the auxiliary regression of the Kelejian-Robinson (KR) test, of order (h_Rx1), Z is the matrix of exogenous variables included in the last regression and $\hat{\gamma}$ the estimation obtained for the corresponding vector of parameters.

As is well known, the asymptotic distribution of the standardised Moran's I is an N(0,1); the two Lagrange multipliers that follow, LM-ERR and LM-EL, have an asymptotic $\chi^2(1)$, the distribution of the KR test is a $\chi^2(m)$, with m being the number of regressors included in the auxiliary regression. The three final tests also have a chi-square distribution, with one degree of freedom in the first two, and two degrees of freedom in the SARMA test.

APPENDIX 2: Systems of regions used in the simulation

R=25

1	*2*	*3*	*4*	*5*
6	*7*	*8*	*9*	*10*
11	*12*	***13***	*14*	*15*
16	*17*	*18*	*19*	*20*
21	*22*	*23*	*24*	***25***

R=100

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
11	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*
21	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*
31	*32*	*33*	*34*	*35*	*36*	*37*	*38*	*39*	*40*
41	*42*	*43*	*44*	***45***	*46*	*47*	*48*	*49*	*50*
51	*52*	*53*	*54*	*55*	*56*	*57*	*58*	*59*	*60*
61	*62*	*63*	*64*	*65*	*66*	*67*	*68*	*69*	*70*
71	*72*	*73*	*74*	*75*	*76*	*77*	*78*	*79*	*80*
81	*82*	*83*	*84*	*85*	*86*	*87*	*88*	*89*	*90*
91	*92*	*93*	*94*	*95*	*96*	*97*	*98*	*99*	***100***

R=225

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*
16	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*
31	*32*	*33*	*34*	*35*	*36*	*37*	*38*	*39*	*40*	*41*	*42*	*43*	*44*	*45*
46	*47*	*48*	*49*	*50*	*51*	*52*	*53*	*54*	*55*	*56*	*57*	*58*	*59*	*60*
61	*62*	*63*	*64*	*65*	*66*	*67*	*68*	*69*	*70*	*71*	*72*	*73*	*74*	*75*
76	*77*	*78*	*79*	*80*	*81*	*82*	*83*	*84*	*85*	*86*	*87*	*88*	*89*	*90*
91	*92*	*93*	*94*	*95*	*96*	*97*	*98*	*99*	*100*	*101*	*102*	*103*	*104*	*105*
106	*107*	*108*	*109*	*110*	*111*	***112***	*113*	*114*	*115*	*116*	*117*	*118*	*119*	*120*
121	*122*	*123*	*124*	*125*	*126*	*127*	*128*	*129*	*130*	*131*	*132*	*133*	*134*	*135*
136	*137*	*138*	*139*	*140*	*141*	*142*	*143*	*144*	*145*	*146*	*147*	*148*	*149*	*150*
151	*152*	*153*	*154*	*155*	*156*	*157*	*158*	*159*	*160*	*161*	*162*	*163*	*164*	*165*
166	*167*	*168*	*169*	*170*	*171*	*172*	*173*	*174*	*175*	*176*	*177*	*178*	*179*	*180*
181	*182*	*183*	*184*	*185*	*186*	*187*	*188*	*189*	*190*	*191*	*192*	*193*	*194*	*195*
196	*197*	*198*	*199*	*200*	*201*	*202*	*203*	*204*	*205*	*206*	*207*	*208*	*209*	*210*
211	*212*	*213*	*214*	*215*	*216*	*217*	*218*	*219*	*220*	*221*	*222*	*223*	*224*	***225***

The contaminated regions appear in bold on grey ground: 13, 45 or 112 in the case of central location or 25, 100 or 225 if it corresponds to a peripheral location.

References

Anselin, L. and R. Florax (1995): Small Sample Properties of Tests for Spatial Dependence in Regression Models. In L. Anselin and R. Florax (eds.): *New Directions in Spatial Econometrics* (pp. 21-74). Berlin: Springer.

Barnett V. and T. Lewis (1994): *Outliers in Statistical Data,3rd Edition.* New York: Wiley.

Belsley, D. (1991): *Conditioning Diagnostics*. New York: Wiley.

Chang, I., G. Tiao and C. Chen (1988): Estimation of Time Series Parameters in the Presence of Outliers. *Technometrics*, 3, 193-204.

Chatterjee, S. and A. Hadi (1988): *Sensitivity Analysis in Linear Regression*. New York: Wiley.

Cressie, N. (1993): *Statistics for Spatial Data*. New York: Wiley.

Florax, R. and T. de Graaff (2004, forthcoming): The Performance of Diagnostic Tests for Spatial Dependence in Linear Regression Models: A Meta-Analysis of Simulation Studies. In L. Anselin, R. Florax and S. Rey (eds.) *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Berlin: Springer

Haining, R. (1994): Diagnostics for Econometric Modeling in Spatial Econometrics. *Journal of Regional Science*, 34, 325-341.

Haining, R. (1995): Data Problems in Spatial Econometric Modeling. In L. Anselin and R. Florax (eds): *New Directions in Spatial Econometrics* (pp 156-171). Berlin: Springer.

Hawkins, D. (1980): *Identification of Outliers*. London: Chapman and Hall.

Kendall, M. and A. Stuart (1977): *The Advanced Theory of Statistics*. London: Charles Griffin & Company Limited, 4th edition.

Lauridsen J. and Mur J. (2004): *Multicollinearity in Cross-Sectional Analysis*. University of Southern Denmark, Department of Economics, Working Paper.

Lauridsen J. and Mur J. (2005): *Multicollinearity and Outliers in Cross-Sectional Analysis*. University of Southern Denmark, Department of Economics, Working Paper.

Peña, D. (1990): Influential Observations in Time Series. *Journal of Business & Economic Statistics*, 8, 235-241.

Perron, P. (1989): The Great Crash, the Oil Price Shock and the Unit Root Hypothesis. *Econometrica*, 57, 1361-1401.

Shekhar, S., C. Lu and P. Zhang (2002): Detecting graph-based spatial outliers. *Intelligent Data Analysis*, 6, 451-468.

Tsay, R. (1988): Outliers, Level Shifts and Variance Changes in Time Series. *Journal of Forecasting*, 7,1-20.

Wartenberg, D. (1989): Exploratory Spatial Analysis: Outliers, Leverage Points and Influence Functions. In D. Griffith (ed.): *Spatial Statistics: Past, Present and Future*. Monograph nº 12 (pp. 133-156). Institute of Mathematical Geography. Syracuse: Syracuse University.

Figure 1: Size and power of Moran's I test under outliers.

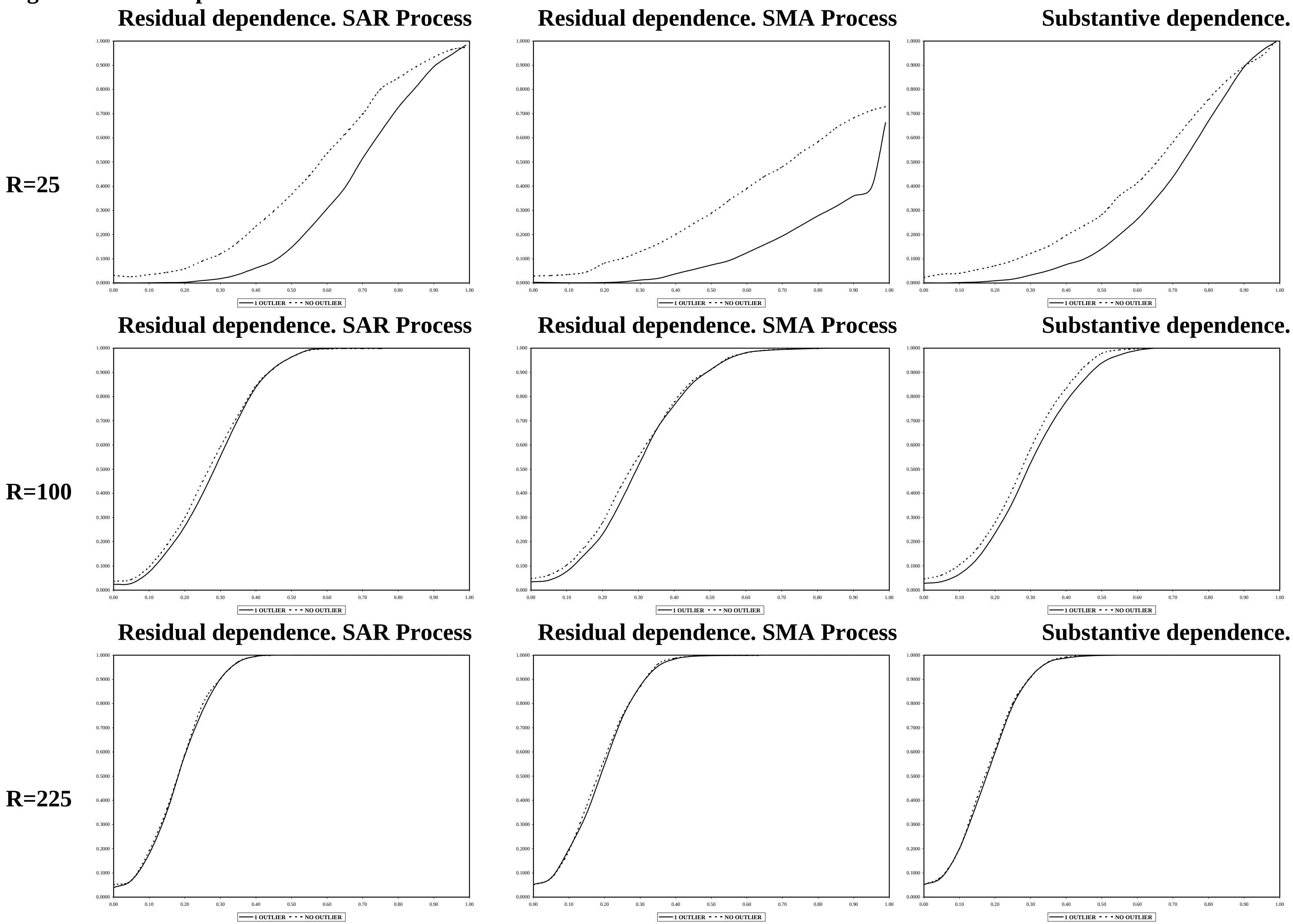

Figure 2: Size and power of LME-ERR test under outliers

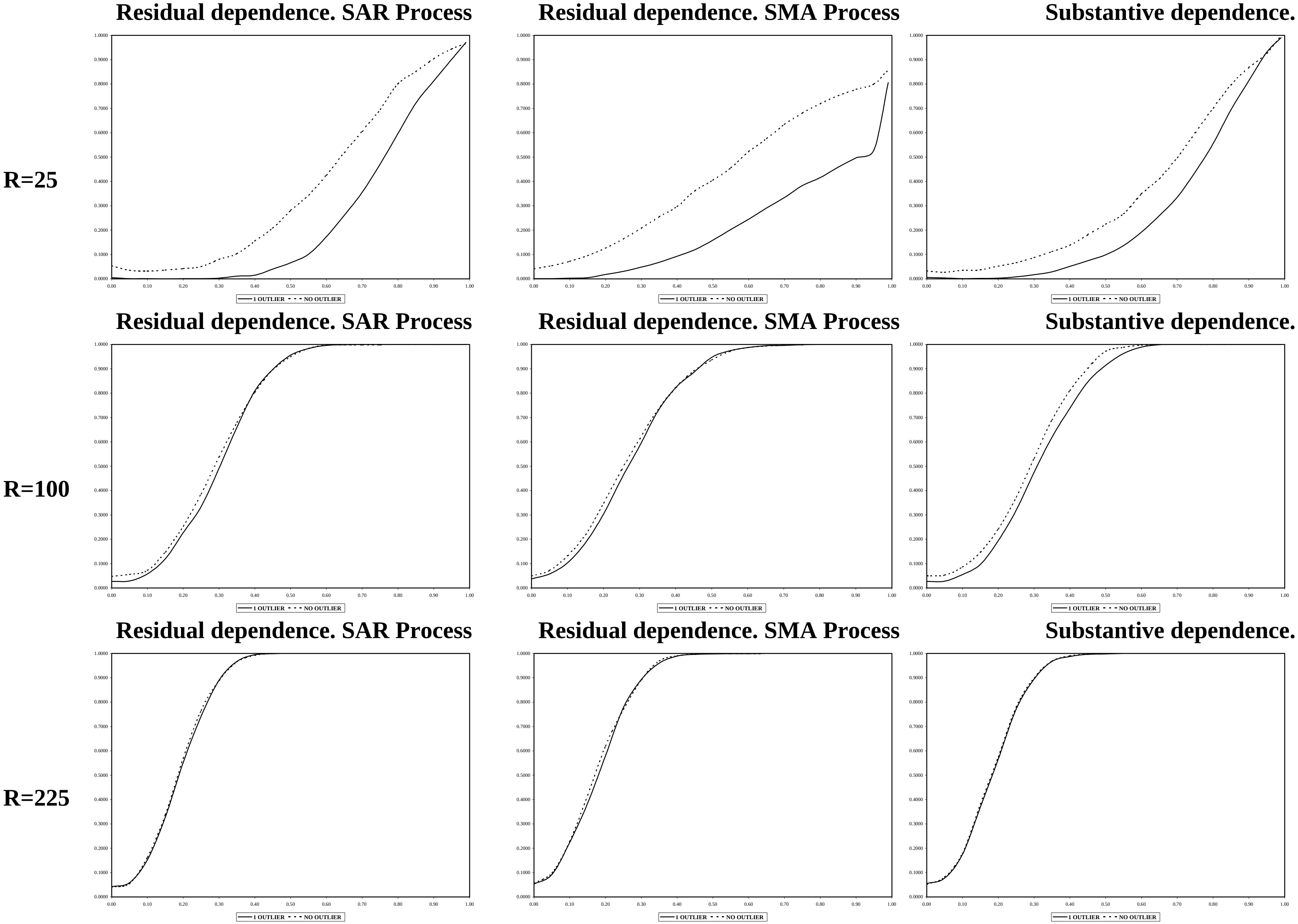

Figure 3: Size and power of KR test under outliers

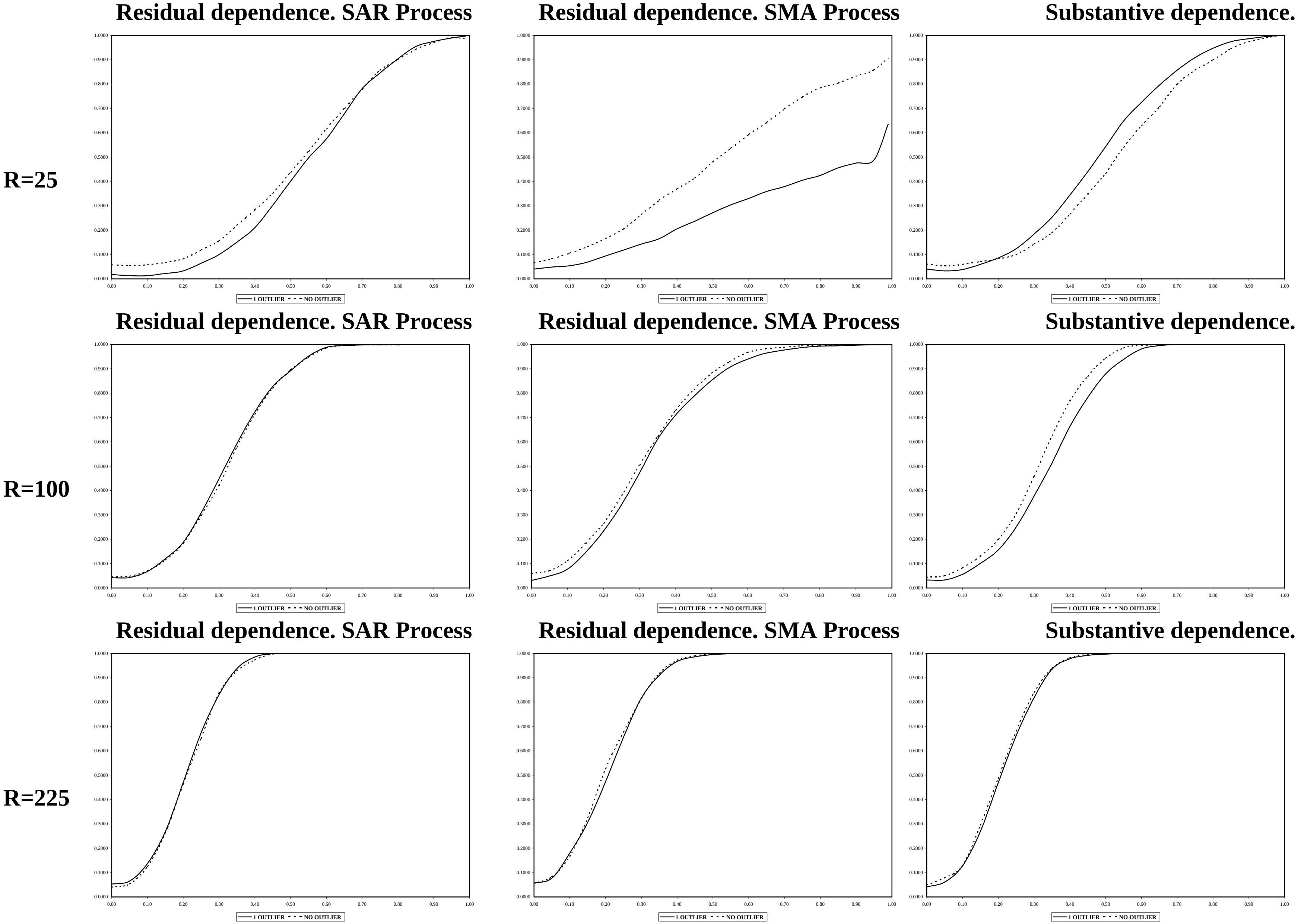

Figure 4: Size and power of LM-EL test under outliers

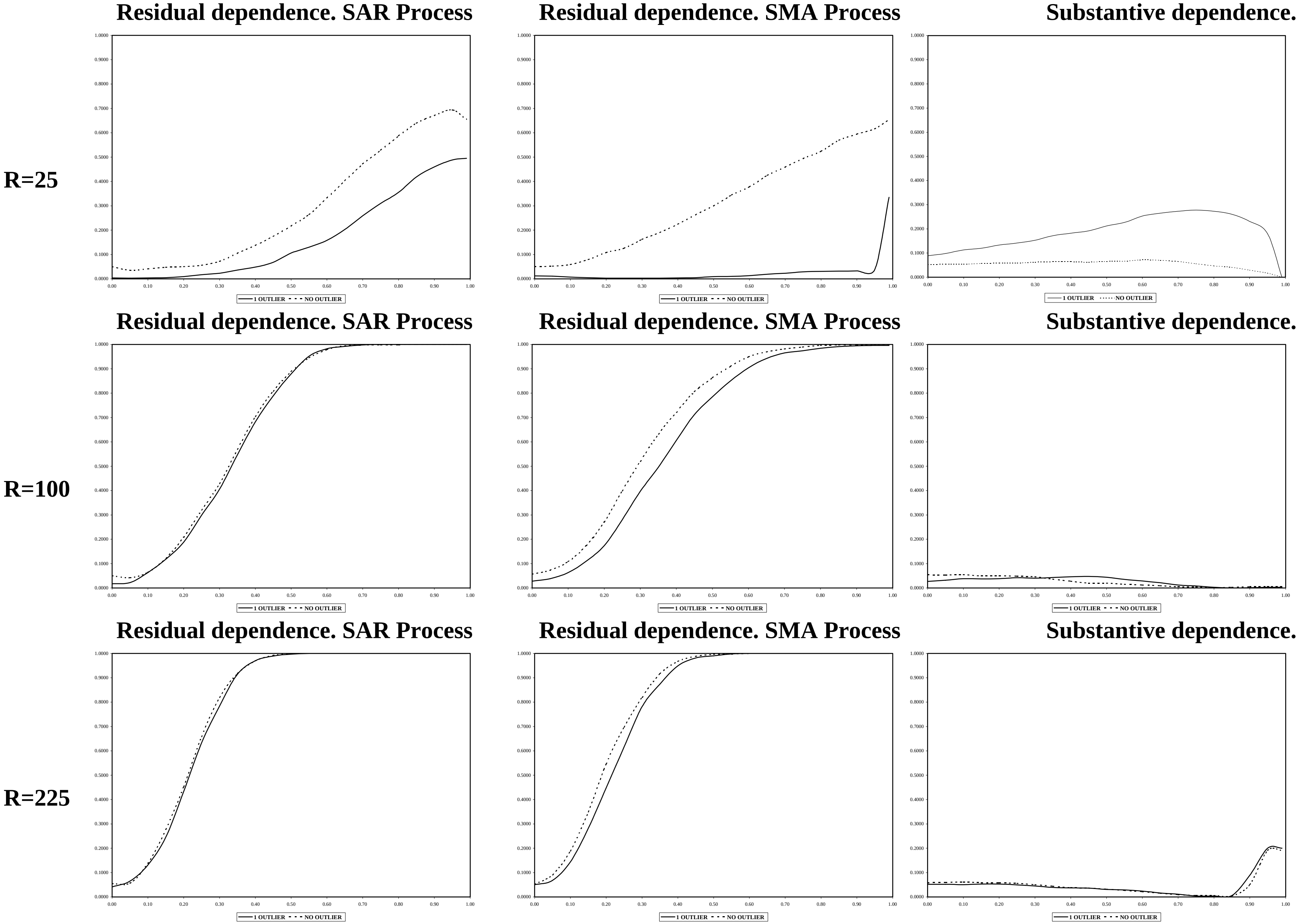

Figure 5: Size and power of LM-LAG test under outliers

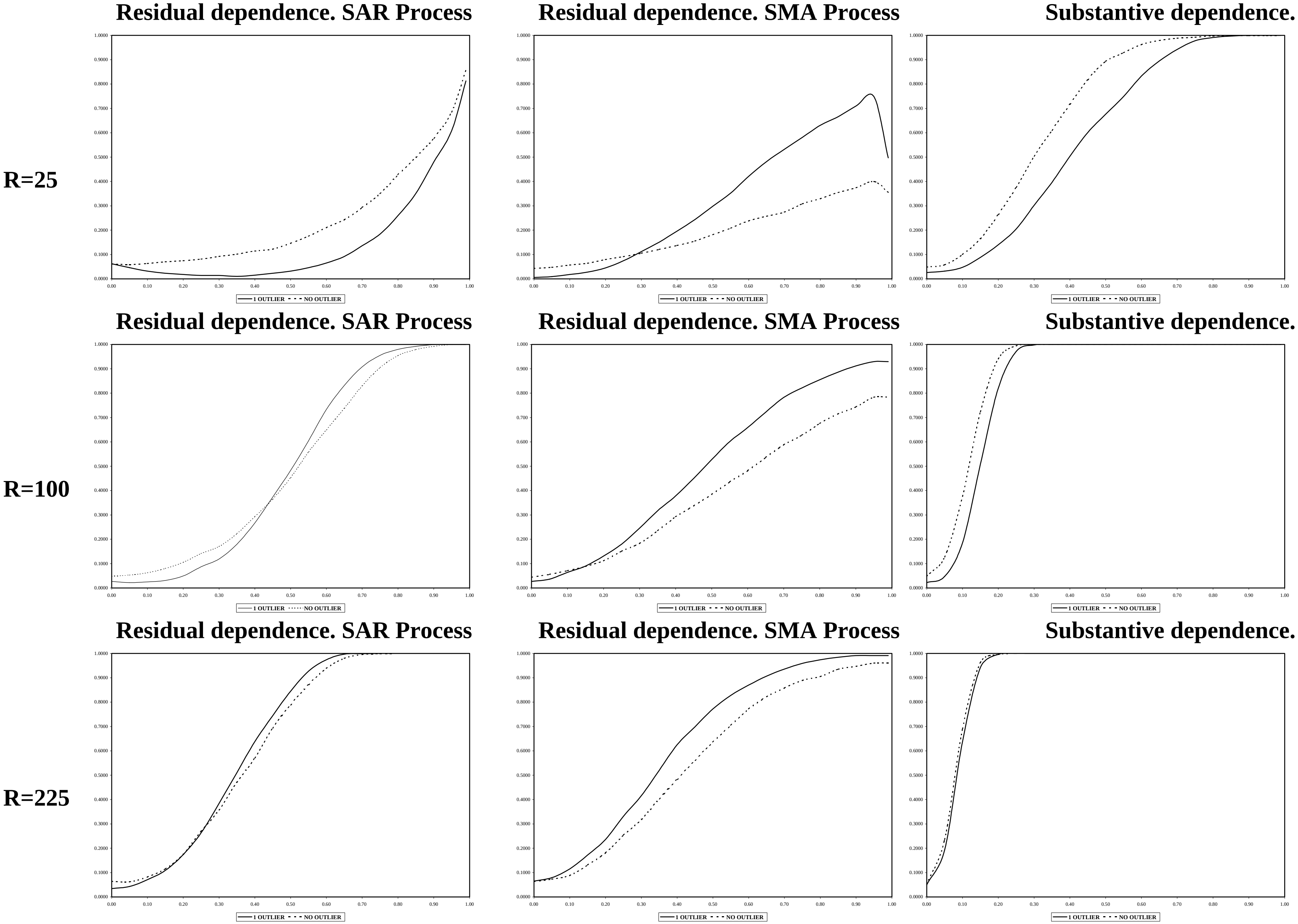

Figure 6: Size and power of LM-LE test under outliers

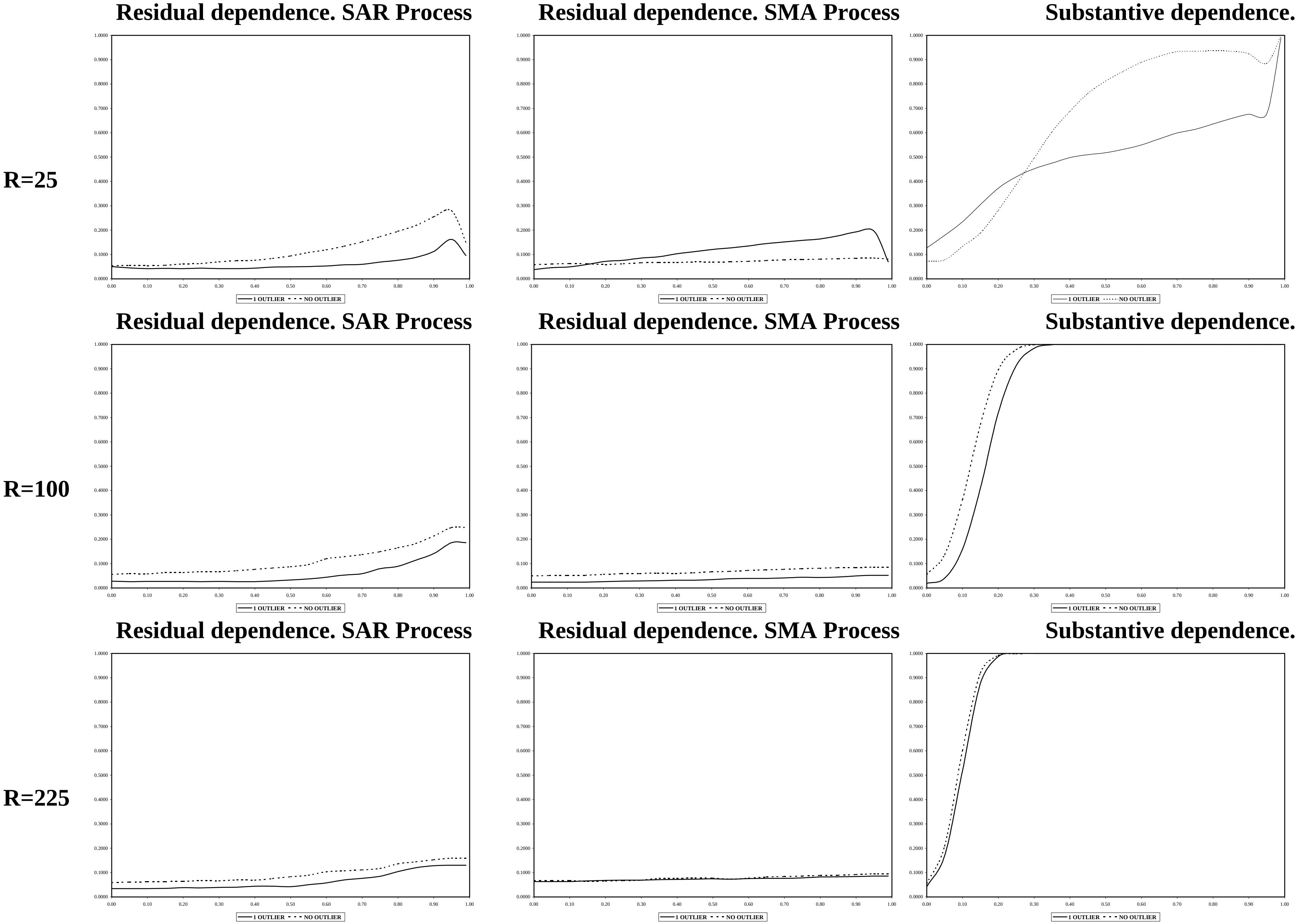

Figure 7: Size and power of SARMA test under outliers